(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: INTELLIGENT PEER-TO-PEER SYSTEM AND METHOD FOR COLLABORATIVE SUGGESTIONS AND PROPAGATION OF MEDIA

(57) Abstract: In a network-based system for recommending media content items based on user preferences, clients contact a server on a periodic basis, independent of the user. In addition to client-server interaction, clients also interact with one another in peer-to-peer fashion. Peers query one another and evaluate their similarity to each other in an interactive comparison of user preferences. When two clients are sufficiently similar, the interaction culminates in the originating client downloading content listings from the targeted peer to generate suggestions for their user. If the two clients are dissimilar, the query may be terminated, or the targeted peer may route the query to a second targeted peer. In addition to the lists of preferences, the originating client may download actual content items from the targeted peers.

INTELLIGENT PEER-TO-PEER SYSTEM AND METHOD FOR COLLABORATIVE SUGGESTIONS AND PROPAGATION OF MEDIA

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates to automated systems and methods for recommending items to users. More particularly, the invention relates to an automated peer-to-peer system and method for collaborative suggestions and propagation of media.

DESCRIPTION OF RELATED ART

The prior art provides various systems for filtering, suggesting and rating of media content items. Common methods of suggesting and rating items occasionally employ collaborative filtering techniques, in which a user's preference profile is compared with profiles of similar users or groups of users. The co-pending application, K. Ali, W. Van Stam, "*Intelligent system and methods of recommending media content items based on user preferences*," U.S. Patent Application Ser. No. 09/xxx,xxx (December 21, 2000) discusses several of these collaborative filtering implementations. In addition, J. Atcheson, J. Miller, "*Method and apparatus for recommending selections based on preferences in a multi-user system*," U.S. Patent No. 5,583,763 (December 10, 1996) describe a system for determining selections that a user is likely to be interested in. A determination is made, based on a user's prior indicated preferences, designated in a preferences list. The list is compared with other users' lists. When a large number of matches is found between two lists, the unmatched entries of the other user's list are extracted. Typically, these implementations require a client-server network environment and a stateful connection between the client and the server. Correlations are calculated on the server, based on data periodically supplied by the client, necessitating monitoring of the client state, thereby raising confidentiality concerns. It would be desirable to provide a collaborative suggestion system in which a stateful connection between client and server is unnecessary, thus reducing concerns about user privacy.

The above-identified co-pending application, K. Ali, *et al.*, *supra*, describes a distributed collaborative filtering engine that guarantees user privacy by eliminating the necessity of

correlating the user to other user's or groups of users. Similarity is calculated on the client side, eliminating the necessity of a stateful connection between the server and the client. The described system, however, employs a client-server architecture in which information is exchanged only between client and server. It would be an advantage to
5   provide a system for collaborative suggestion in a peer-to-peer environment, which makes opportunistic use of an existing network connection, wherein peers evaluate their similarity to one another.

Peer-to-peer file sharing systems are becoming increasingly common. For example, the
10   "Gnutella Support Pages," http:// gnutella.wego.com (no date) describe a peer-to-peer network composed of a multiplicity of Gnutella clients, in which the client software includes an integrated search engine and file server. The Gnutella network changes constantly, according to the number of Gnutella clients that are on the network at any given time. No server exists, and the network infrastructure is provided by a publicly
15   accessible network, such as the Internet. In order to access the Gnutella network, a user must have the network address of at least one other Gnutella client that is currently connected. A user in search of a particular information object, a digital music file, or a recipe, for example, may send a query over the network. The query is passed from client to client until the object is located or the query is terminated. While the Gnutella
20   client allows the creation of a dynamic peer-to-peer network, and sharing of files between clients, the query process is user-initiated: queries are formulated and launched by the user with no automation of the query process. Additionally, the Gnutella network is primarily directed to file sharing, in which media content items are shared or propagated between users. There is no capability of comparing user profiles between
25   clients in order to generate collaborative suggestions. Furthermore, the Gnutella network is concerned exclusively with the peer-to-peer network paradigm.

It would be a technological advance to provide a system for collaborative suggestions and media propagation that did not require a stateful connection between a client and
30   server, thus safeguarding privacy of individual users. It would be a great advantage to implement such a system as a peer-to-peer based system that was capable of operating in parallel with client-server based suggestion systems, opportunistically employing the same network connection, wherein suggestions generated by both systems are presented in the same software interface. Furthermore, it would be
35   desirable to automate the peer-to-peer system, so that clients could initiate and carry out interactions with each other without direction or intervention by a user.

5   ## SUMMARY OF THE INVENTION

A network-based intelligent system for predicting ratings for items of media content according to how likely they are to appeal to a user provides a parallel, peer-to-peer system and method for collaborative suggestions and propagation of media. Using a
10   typical client-server architecture, clients contact a server on a periodic basis, independent of the user. In addition to client-server interaction, clients also make opportunistic use of the network connection to interact with one another in peer-to-peer fashion. The server organizes clients into groups and provides each client within a group with the network address of all other clients in the group. An originating client queries a targeted peer b y
15   transmitting a list indicative of its user's preferences. The targeted peer evaluates the similarity of the transmitted list with a list of its own. If the two clients are sufficiently similar, the comparison continues in an interactive fashion. After the initial determination by the targeted peer, the exchange of information proceeds in a stepwise manner, with the originating client evaluating similarity at each stage. If the two clients are dissimilar, either
20   the originating client or the targeted peer may terminate the query, depending on the stage of the interaction; or the targeted peer may route the query to a second targeted peer. The interaction culminates in the originating client downloading content listings from the targeted peer to generate suggestions for their user. In addition to the lists of preferences, the originating client may download actual content items from the targeted
25   peers.

BRIEF DESCRIPTION OF THE DRAWINGS

30   Figure 1 provides a block diagram of a network-based video recording system, according to the invention; and

Figure 2 provides a block diagram illustrating peer-to-peer interaction among the clients of the system of Figure 1, according to the invention.
35

DETAILED DESCRIPTION

The co-pending application, K. Ali, *et al.*, *supra*, provides an intelligent, distributed system for recommending items of content to a user, based on the user's expressed

preferences. The described system, shown here in Figure 1, employs a client-server network architecture in which each of a plurality of clients is periodically in contact with the server. In the described system, all interaction occurs between the client 11 and the server 10, as Figure 1 shows. However, the current invention extends the suggestion

5   generating capability of the previous system by making opportunistic use of the network connection to provide peer-to-peer interaction among the clients, so that the knowledge embodied on each of the clients is further leveraged by exchanging suggestions and content directly between clients, in peer-to-peer fashion. In the parallel peer-to-peer system, shown in Figure 2, the server 10 is operative to organize

10   simultaneously connected clients 11 into one or more groups 20, whereupon the server provides each client within a group 20 the network address of all other clients in the same group. Beyond this initial function, all interaction occurs among the clients, independently of the server, in parallel with the client-server interaction. In the system of the parent application, the client is in contact with the server for a short period of time to download

15   current program guide data and to exchange data with the server for the purpose of generating collaborative suggestions. Since the time of connection varies from day-to-day, in a large community of users, the selection of clients connected to the network at any given time is apt to be fairly random, so that, over time, each client is randomly exposed to a large population of other clients, thereby enabling a rich exchange of

20   information.

While the preferred embodiment of the invention employs the Internet as its network infrastructure, other publicly accessible telecommunications networks would also be suitable: for example, a cable television network. The presently preferred embodiment

25   of the invention employs a dial up network connection. However, hardwired connections, for example, coaxial or fiber optic cable would also be suitable, particularly in the case of broadband implementations of the invention. Furthermore, wireless connections would also be consistent with the spirit and scope of the invention. The type of media involved is highly variable. While the present embodiment of the invention is concerned

30   primarily with various types of television programming, the invention also finds application with text files, web sites, books, digital music; in short, almost any type of digital media.

Communication occurs directly between the clients, in a manner that is difficult to trace or

35   monitor. Furthermore, network addresses are dynamically assigned, and are valid only for the duration of the connection. Thus, the temporary and anonymous nature of the peer-to-peer interaction provides an important safeguard to user privacy.

During their interaction, two clients go through an interactive comparison procedure, in which they compare information that is highly indicative of their respective user's interests. At each step of the procedure, correlation, or similarity is calculated. If the correlations converge, the originating client may request user preference lists from the target peer. If such a convergence isn't reached or the correlation is deemed too low, then the query can be terminated or relayed to a second targeted peer. Various embodiments of the method of interaction between peers are described below.

PEER-TO-PEER SUGGESTION SYSTEM

In the invented system, a pool of client devices is simultaneously connected to a network. The presently preferred embodiment of the invention incorporates clients of a distributed personal video recording system, in which every client is a dedicated video recording unit. Resident on every client are several lists that are highly indicative of a user's preferences. Minimally, the lists include:

- Lists of recorded items that are currently available for viewing;
- Lists of various types of items to record (single programs, series, programs of a particular subject or type, programs having a particular actor, and so on); and
- Lists of rated items.

I. PEER-TO-PEER INTERCONNECTION

As previously indicated, clients connect to a central server periodically. In a larger community of users, there is a constant group of clients connected to the server, dynamic in nature, in which clients constantly come and go.

The server organizes clients by connection time, so that clients that have connected recently are all placed in the same group; additionally, the server controls group size. A larger group size is preferable in order to maximize the opportunity for interaction among clients. The server provides each client in the group network with network addresses to all other clients in the group. Following group formation, each client is free to contact any other client in the group directly.

II. PEER-TO-PEER INTERACTION

The lists mentioned above may be hundreds or even thousands of items in length. Thus, exchanging an entire list or set of lists between clients may be impractical, particularly in the case of a narrowband dial up network connection. Therefore, a variety of methods have been provided to facilitate peer-to-peer interaction in a manner that economizes on time and network resources.

Method 1:   Small lists having dense meaning.

Each client has one or more compact lists of items that are highly representative of the user's preferences. These may be the list of series to record, or the list of items currently available. Each of them armed with such a list, an originating client (A) and targeted client
5   (B) interact as follows:

- Client A sends a query to B, consisting of the list, and the size of the other lists it has to exchange.

10
- Client B receives the query and compares A's list with its own corresponding list to determine the number of elements it shares in the list with client A. It now computes the similarity between A and B, according to an algorithm, resident on both A and B, in which:

    Similarity = Number of elements in common between A and B /
15   number of elements reported by A. Similarity is expressed as a value in a range of approximately 0 to 1. Client B may respond in one of several ways:

If the similarity value is less than a first predetermined threshold, it can respond to A that B has nothing that A might be interested in.

20

If the similarity value exceeds the predetermined threshold, B can respond that A should be interested in B. In addition, B sends more information with the response that indicates relative sizes of B's lists, compared to A's. For example, B might say "Yes, you are 80% interested in me, and I have twice the number of ratings data that you
25   have, and a third your size of recording history". Client A's response is described further below.

If the similarity value is less than the predetermined threshold, B may pass the query from A to a second targeted peer. Upon redirecting, a reference to B is appended to
30   the query so that successive peers know not to pass the query to B. It also allows A to know which clients have evaluated its query when it receives a response. In any further queries originated by A, it will then skip over those that have already been queried.

If Client A receives a positive response from a targeted peer, comprising similarity
35   values, relative list sizes and a listing of peers that have evaluated the query, further interaction is determined by A. If A determines that the similarity value is less than a second predetermined threshold, A terminates the query. Having terminated the query, A may direct additional queries to other clients within the group that have not yet been queried. If the similarity value exceeds the second threshold, client A evaluates which

lists it wants to retrieve from the responding peers. If the relative size of any of the other lists is sufficient, A may request the complete list from the responding peer. For example, A may request a full recording history from one respondent and a complete list of ratings from another. Upon receiving the lists, A further evaluates the lists for correlation, and uses them to generate collaborative suggestions for the user. Appropriate methods for computing correlation and generating collaborative suggestions are described in the co-pending application previously mentioned, K. Ali, *et al.*, *supra*.

Method 2: Iterative disclosure.

The previous method assumed that sending a full list of items can be done in small packets that travel easily over the connecting network. When the lists are large, like the recording history, or the ratings list, this method is inefficient. An alternate approach is the sharing of the lists in successive blocks, in iterative fashion. In the following description, as above, Client A is the originator and Client B is the target:

- Client A takes a block that includes the first n consecutive items in it's list and transmits it to B.

- B receives the list and determines matching items; and transmits a list of matching items to A.

- A creates a temporary matching items list, where it saves the list received from B, and transmits another block of n items to B. B responds with another list of matching items, which A adds to its temporary list. With each succeeding response from B, A evaluates what portion of the total number of items shared with B the two peers have in common. When enough items have been tested by repeating the above steps, the ratio remains relatively constant from one iteration to the next. At that point, the similarity value for the partial data is representative of the similarity value for the entire list. The query then proceeds as described above, in which A requests from the target the lists it wishes to receive.

Method 3: Iterative disclosure of ratings
Similar to the previous method:

- Again, A Sends partial lists to B, and B echoes back what it has. Meanwhile A and B keep track of what they share. In addition, the lists that A sends to B contain A's ratings of the items, and B's answers contain B's ratings of the items matching items. Thus, both A and B maintain a list of items, with both A's and B's ratings. This allows A and B to use standard correlation math to determine how well A and B. As the size of the list of items they have in common grows, the

correlation factors will tend to mirror the real correlation factor that would be obtained if all available data was known to both A and B.

- According to a minimum confidence requirement standard (based on the number of items shared from A, and number of items that are shared) the process is repeated until the standard is satisfied. After that, A and B can determine if the correlation is high enough to proceed with sharing full lists or to terminate the connection, in the event of an insufficient correlation.

Those skilled in the art will recognize that the roles of originator and target have been employed for purposes of description. In actual fact, each client within a group is both an originator and a target, wherein each client directs queries to other clients and receives queries from other clients. Furthermore, while the client-server system of the copending application and the present peer-to-peer system have been treated as parallel systems for the sake of description, the skilled practitioner will appreciate that they are, in fact, one system having parallel functionalities. Thus, the client units interact with the server in the usual client-server fashion, and simultaneously also have the peer-to-peer functionality herein described. The invention is implemented using conventional techniques well known to those skilled in the art of software engineering and computer programming. The network implementation will be apparent to those skilled in the design and administration of data networks.

While a narrow band dial up connection renders the transmission between clients of large media files impractical, broadband implementations of the invention, wherein connection is by means of fiber optic or coaxial cable, DSL, T1 or T3, or the like, enable the peer-to-peer sharing of the underlying content, as well as the user preference files.

Although the invention has been described herein with reference to certain preferred embodiments, one skilled in the art will readily appreciate that other applications may be substituted without departing from the spirit and scope of the present invention. Accordingly, the invention should only be limited by the Claims included below.

CLAIMS

What is claimed is:

1.   A peer-to-peer system for collaborative suggestions and media propagation comprising:

a network, said network comprising a plurality of clients and at least one server;

one or more lists for each client indicative of interests of a corresponding user;

means for estimating similarity between said clients, based on said lists; and

means for automatically querying one client by another; so that similar clients may share information;

wherein said peer-to-peer system operates in parallel with a client-sever system, making opportunistic use of an existing network connection.

2.   The system of Claim 1,wherein each of said clients is periodically in contact with said server, and wherein at least a portion of said clients are simultaneously in contact with said server.

3.   The system of Claim 2, wherein said server organizes said clients simultaneously in contact with said server into one or more groups, and wherein said means for automatically querying one client by another comprises a network address for each client, wherein said server transmits the network address for each client within a group to all other clients within said group, so that clients within a group may query each other directly.

4.   The system of Claim 3, wherein said one or more lists include at least one of:

a list of items of media content stored on the respective client;

a list of items desired by said user; and

a list of items rated by said user.

5.   The system of Claim 4, wherein an originating client queries a first targeted peer within said group, and wherein said originating client transmits a list of items indicative of said user's interests.

6.   The system of Claim 5, wherein said means for estimating similarity between said clients comprises an algorithm, said algorithm residing on each of said clients, and wherein said targeted peer compares said transmitted list with a list of its own and estimates similarity between the two lists according to said algorithm, said similarity
5   expressed as a value.

7.   The system of Claim 6, wherein said algorithm includes the steps of:
       determining the number of items common to both abbreviated lists; and
       dividing the number of common items by the total number of items on the list
10   from the originating client, wherein said value results.

8.   The system of Claim 7, wherein said value is in a range of approximately 0 to 1.

9.   The system of Claim 6, wherein the targeted peer responds by any of:
15   for a similarity estimate falling below a predetermined threshold value, said targeted peer returns said query without providing additional information, while adding a reference to said query so that said originating client doesn't direct further queries to said first targeted peer;
            for a similarity estimate falling below said predetermined threshold value, said first
20   targeted peer redirects said query to a second targeted peer, wherein said first targeted peer adds a reference to said query so that subsequent peers receiving the query don't redirect the query to the first targeted peer, and so that said originating client has a record of peers already queried; and
            for a similarity estimate equaling or exceeding said predetermined threshold
25   value, said first targeted peer responds to said originating client by transmitting the similarity estimate along with sizes of complete lists relative to complete list sizes on said originating client.

10.   The system of Claim 9, wherein said originating client evaluates similarity
30   estimates, and relative list sizes received from responding targeted peers and responds by any of:
       terminating said query;
       directing said query to peers that have not yet been queried;
       requesting one or more lists from one or more of said responding targeted
35   peers.

11.   The system of Claim 10, wherein said originating client evaluates lists received from said responding targeted peers and utilizes unique items from said lists to generate suggestions and recommendations for said user.

12.   The system of Claim 11, wherein said originating client requests items of media content from said responding peers, based on contents of said lists.

5   13.   The system of Claim 11, wherein said lists are transmitted in discrete packets.

14.   The system of Claim 5, wherein said originating client iteratively queries said targeted peer, wherein an abbreviated list is transmitted with each query, said abbreviated lists comprising successive blocks of n consecutive entries from a complete
10   list.

15.   The system of Claim 14, wherein said targeted peer iteratively evaluates each abbreviated list received for items common to both originating client and targeted peer and transmits a list of said common items to said originating client.
15

16.   The system of Claim 15, wherein said originating client evaluates said lists of common items received from said targeted peer and estimates similarity for the corresponding abbreviated list, and wherein an average of said similarity estimates constitutes an estimate of similarity between said originating client and said targeted
20   peer.

17.   The system of Claim 16, wherein said lists exchanged between said originating client and said targeted peer also include user ratings for each item, and wherein a correlation is calculated for each abbreviated list based on said user ratings.
25

18.   The system of Claim 17, wherein correlations derived from said abbreviated lists converges, as the number of common items grows, to an actual correlation that would be obtained if all available data from said originating server and said targeted peer were known.
30

19.   The system of Claim 18, wherein said iterative queries continue until confidence in said calculated correlation equals or exceeds a minimum confidence level, based on total number of items shared from said originating server and total number of common items.

35   20.   The system of Claim 19, wherein said originating client:
            terminates said query based on said calculated correlation; or
            requests one or more complete lists from said targeted client, based on said calculated correlation.

21.   The system of Claim 1, wherein said system is a component of a network-based personal video recording system.

22.   A peer-to-peer method of generating collaborative suggestions and propagating items of media content comprising the steps of:

    providing a network, said network comprising a plurality of clients and at least one server;

    providing one or more lists for each client indicative of interests of a corresponding user;

    estimating similarity between said clients, based on said lists; and

    automatically querying one client by another; so that similar clients may share information, wherein said peer-to-peer method executes in parallel with a client-server method, opportunistically using an existing network connection.

23.   The method of Claim 22, wherein each of said clients is periodically in contact with said server, and wherein at least a portion of said clients are simultaneously in contact with said server.

24.   The method of Claim 22, further comprising the steps of:

    organizing said clients simultaneously in contact with said server into one or more groups; and

    transmitting a network address for each client within a group to all other clients within said group by said server, so that clients within a group may query each other directly.

25.   The method of Claim 24, wherein said one or more lists include at least one of:

    a list of items of media content stored on the respective client;

    a list of items desired by said user; and

    a list of items rated by said user.

26.   The method of Claim 24, wherein said step of automatically querying one client by another comprises:

    querying a first targeted peer within said group by an originating client, wherein said originating client transmits a list of items indicative of said user's interests.

27.   The method of Claim 26, wherein said step of estimating similarity between said clients comprises the steps of:

    comparing said transmitted list with a corresponding list of its own by said targeted peer; and

estimating similarity between the two lists according to an algorithm, said algorithm residing on each of said clients, said similarity expressed as a value.

28.   The method of Claim 27, wherein said algorithm includes the steps of:
      determining the number of items common to both lists; and
      dividing the number of common items by the total number of items on the abbreviated list from the originating client, wherein said value results.

29.   The method of Claim 28, wherein said value is in a range of approximately 0 to 1.

30.   The method of Claim 28, further comprising any of the steps of:
      for a similarity estimate falling below a predetermined threshold value, returning said query by said first targeted peer without providing additional information, while adding a reference to said query so that said originating client doesn't direct further queries to said first targeted peer;
      for a similarity estimate falling below said predetermined threshold value, redirecting said query by said first targeted peer to a second targeted peer, wherein said first targeted peer adds a reference to said query so that subsequent peers receiving the query don't redirect the query to the first targeted peer, and so that said originating client has a record of peers already queried; and
      for a similarity estimate equaling or exceeding said threshold minimum value, responding to said originating client by said first targeted peer by transmitting the similarity estimate along with sizes of complete lists relative to complete list sizes on said originating client.

31.   The method of Claim 30, further comprising the step of:
      evaluating similarity estimates and relative list sizes, by said originating client, received from responding targeted peers

32.   The method of Claim 31, further comprising one of the steps of:
      terminating said query by said originating client;
      directing said query to peers that have not yet been queried; and
      requesting one or more lists from one or more of said responding targeted peers.

33.   The method of Claim 32 further comprising the step of:
      utilizing unique items from said lists to generate suggestions and recommendations for said user by said originating client.

34.   The method of Claim 32, further comprising the step of:

by said originating client, requesting items of media content from said responding peers, based on contents of said lists.

35.   The method of Claim 32, wherein said lists are transmitted in discrete packets.

36.   The method of Claim 24, wherein said step of automatically querying one client by another comprises:

an originating client iteratively querying a targeted peer, wherein an abbreviated list is transmitted with each query, said abbreviated lists comprising consecutive blocks of n consecutive entries from a complete list, starting at a beginning of said complete list.

37.   The method of Claim 36, further comprising the steps of

said targeted peer iteratively evaluating each abbreviated list received for items common to both originating client and targeted peer; and

and transmitting a list of said common items to said originating client.

38.   The method of Claim 37, wherein said step of estimating similarity between clients comprises the steps of:

said originating client evaluating said lists of common items received from said targeted peer;

and estimating similarity for the corresponding abbreviated list, wherein an average of said similarity estimates constitutes an estimate of similarity between said originating client and said targeted peer.

39.   The method of Claim 38, further comprising the step of:

calculating a correlation for each abbreviated list based on user ratings for each item, wherein said lists exchanged between said originating client and said targeted peer also include said user ratings.
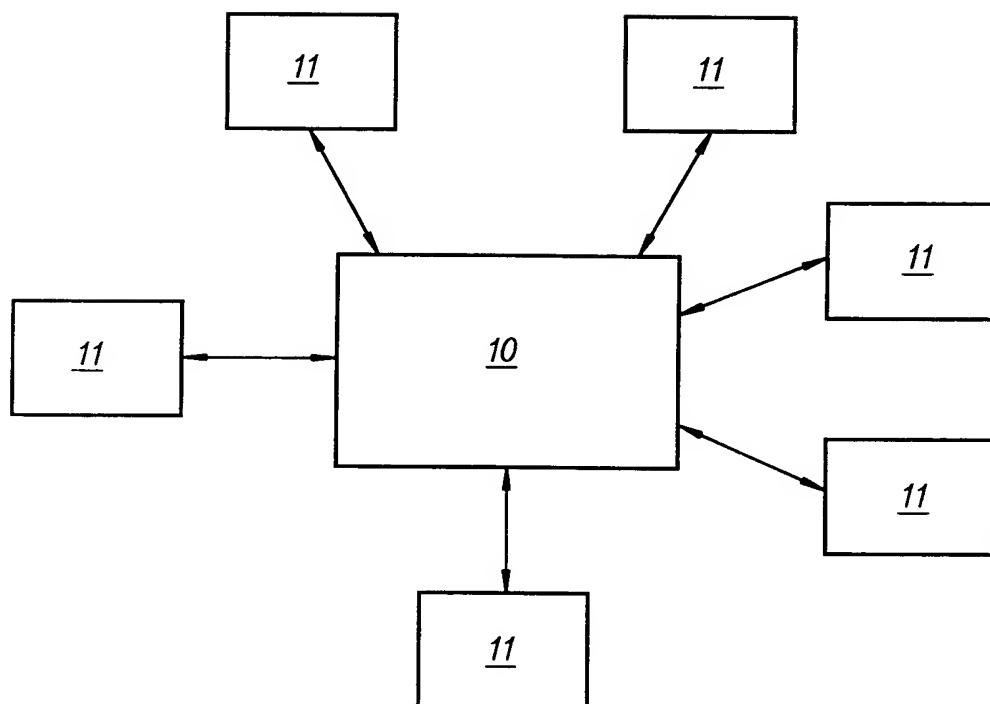
40.   The method of Claim 39, wherein correlations derived from said abbreviated lists converge, as the number of common items grows, to an actual correlation that would be obtained if all available data from said originating server and said targeted peer were known.

41.   The system of Claim 39, wherein said iterative queries continue until confidence in said calculated correlation equals or exceeds a minimum confidence level, based on total number of items shared from said originating server and total number of common items.

42.   The method of Claim 41, further comprising one of the steps of:
        said originating client terminating said query based on said calculated correlation; and
        said originating client requesting one or more complete lists from said targeted client, based on said calculated correlation.
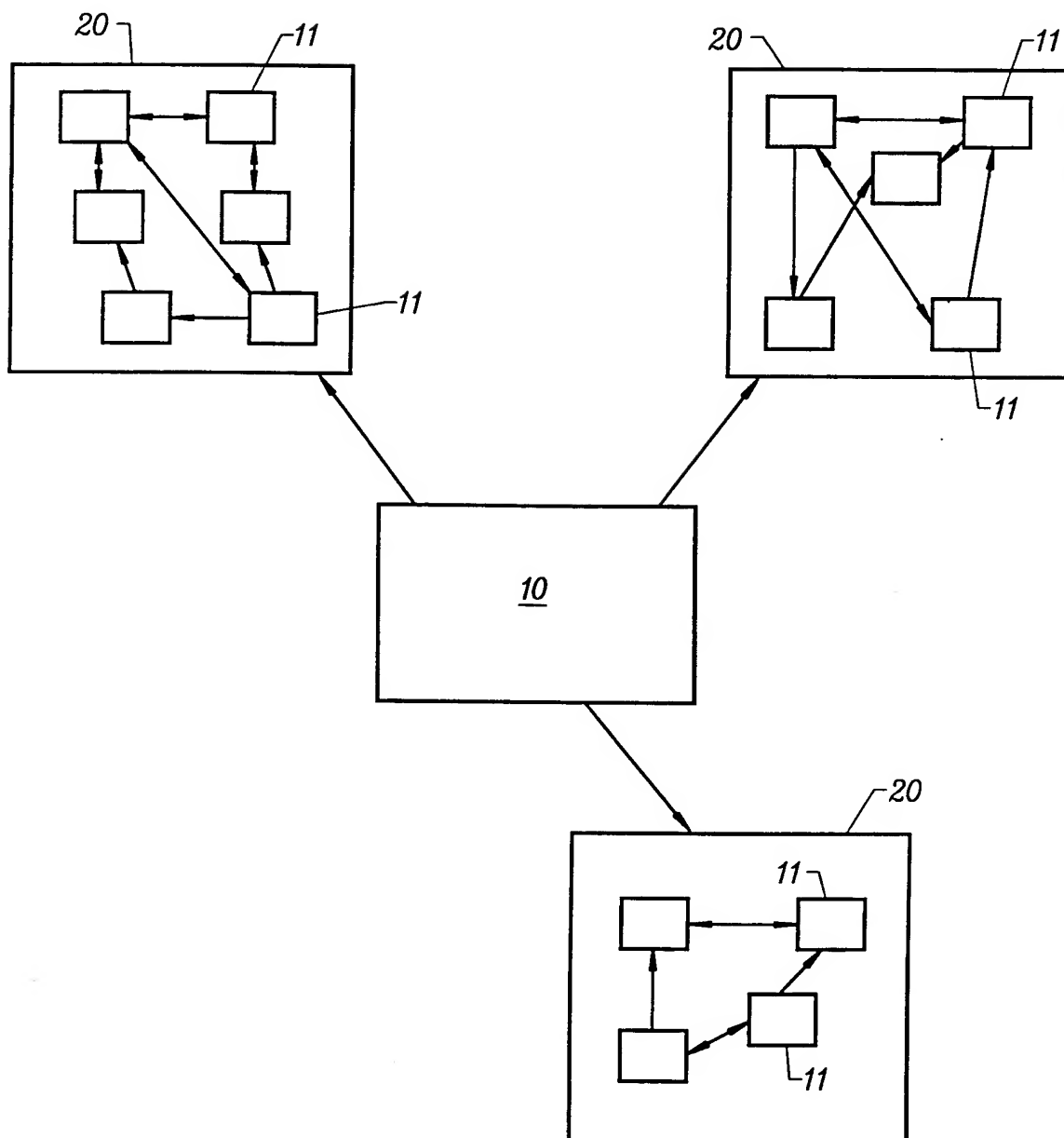
FIG. 1

FIG. 2